

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : <b>C07K 14/435, C12N 1/00, 1/15, 1/21, 5/10, 15/12, 15/63</b>		A1	(11) International Publication Number: <b>WO 00/34325</b> (43) International Publication Date: <b>15 June 2000 (15.06.00)</b>
(21) International Application Number: <b>PCT/US99/29472</b> (22) International Filing Date: <b>10 December 1999 (10.12.99)</b>		(81) Designated States: JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(30) Priority Data: 09/210,330 11 December 1998 (11.12.98) US 09/457,556 9 December 1999 (09.12.99) US		<b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(71) Applicant: CLONTECH LABORATORIES, INC. [US/US]; 1020 East Meadow Drive, Palo Alto, CA 94303 (US).			
(72) Inventors: LUKYANOY, Sergey Anatolievich; ul. Golubinskaya 13/1-161, Moscow (RU). FRADKOV, Arcady Fedorovich; ul. Dnepropetrovskaya, 35/2-14, Moscow, 113570 (RU). LABAS, Yulii Aleksandrovich; ul. Generala Tyuleneva, 35/416, Moscow, 117465 (RU). MATZ, Mikhail Vladimirovich; ul. Teplii stan, 7/2-28, Moscow, 117465 (RU). FANG, Yu; 583 Enos Street, Fremont, CA 94539 (US). TAN, Wenyan; 1220 North Bascom Avenue, #37, San Jose, CA 95128 (US).			
(74) Agent: ADLER, Benjamin, A.; McGregor & Adler, 8011 Candle Ln., Houston, TX 77071 (US).			
(54) Title: FLUORESCENT PROTEINS FROM NON-BIOLUMINESCENT SPECIES OF CLASS ANTHOZOA, GENES ENCODING SUCH PROTEINS AND USES THEREOF			
(57) Abstract			
The present invention is directed to novel fluorescent proteins from non-bioluminescent organisms from the Class Anthozoa. Also disclosed are cDNAs encoding the fluorescent proteins.			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	MW	Malawi	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Kazakhstan	PT	Portugal		
CU	Cuba	LC	Saint Lucia	RO	Romania		
CZ	Czech Republic	LI	Liechtenstein	RU	Russian Federation		
DE	Germany	LK	Sri Lanka	SD	Sudan		
DK	Denmark	LR	Liberia	SE	Sweden		
EE	Estonia			SG	Singapore		

**FLUORESCENT PROTEINS FROM NON-BIOLUMINESCENT SPECIES  
OF CLASS ANTHOZOA, GENES ENCODING SUCH PROTEINS AND  
5 USES THEREOF**

10

**BACKGROUND OF THE INVENTION**

Cross-reference to Related Application

This is a divisional application of U.S.S.N. 09/210,330 filed  
15 on December 11, 1998.

Field of the Invention

This invention relates to the field of molecular biology.  
More specifically, this invention relates to novel fluorescent proteins,  
20 cDNAs encoding the proteins and uses thereof.

Description of the Related Art

Fluorescence labeling is a particularly useful tool for marking a protein, cell, or organism of interest. Traditionally, a  
25 protein of interest is purified, then covalently conjugated to a fluorophore derivative. For *in vivo* studies, the protein-dye complex is then inserted into cells of interest using micropipetting or a method of reversible permeabilization. The dye attachment and insertion steps, however, make the process laborious and difficult to control. An

alternative method of labeling proteins of interest is to concatenate or fuse the gene expressing the protein of interest to a gene expressing a marker, then express the fusion product. Typical markers for this method of protein labeling include  $\beta$ -galactosidase, firefly luciferase and bacterial luciferase. These markers, however, require exogenous substrates or cofactors and are therefore of limited use for *in vivo* studies.

A marker that does not require an exogenous cofactor or substrate is the green fluorescent protein (GFP) of the jellyfish 10 *Aequorea victoria*, a protein with an excitation maximum at 395 nm, a second excitation peak at 475 nm and an emission maximum at 510 nm. GFP is a 238-amino acid protein, with amino acids 65-67 involved in the formation of the chromophore.

Uses of GFP for the study of gene expression and protein 15 localization are discussed in detail by Chalfie et al. in *Science* 263 (1994), 802-805, and Heim et al. in *Proc. Nat. Acad. Sci.* 91 (1994), 12501-12504. Additionally, Rizzuto et al. in *Curr. Biology* 5 (1995), 635-642, discuss the use of wild-type GFP as a tool for visualizing subcellular organelles in cells, while Kaether and Gerdes in *Febs Letters* 20 369 (1995), 267-271, report the visualization of protein transport along the secretory pathway using wild-type GFP. The expression of GFP in plant cells is discussed by Hu and Cheng in *Febs Letters* 369 (1995), 331-334, while GFP expression in *Drosophila* embryos is described by Davis et al. in *Dev. Biology* 170 (1995), 726-729.

25 Crystallographic structures of wild-type GFP and the mutant GFP S65T reveal that the GFP tertiary structure resembles a barrel (Ormö et al., *Science* 273 (1996), 1392-1395; Yang, et al., *Nature Biotechnol* 14 (1996), 1246-1251). The barrel consists of beta sheets in a compact structure, where, in the center, an alpha helix containing

the chromophore is shielded by the barrel. The compact structure makes GFP very stable under diverse and/or harsh conditions such as protease treatment, making GFP an extremely useful reporter in general. However, the stability of GFP makes it sub-optimal for 5 determining short-term or repetitive events.

A great deal of research is being performed to improve the properties of GFP and to produce GFP reagents useful and optimized for a variety of research purposes. New versions of GFP have been developed, such as a "humanized" GFP DNA, the protein product of 10 which has increased synthesis in mammalian cells (Haas, et al., *Current Biology* 6 (1996), 315-324; Yang, et al., *Nucleic Acids Research* 24 (1996), 4592-4593). One such humanized protein is "enhanced green 15 fluorescent protein" (EGFP). Other mutations to GFP have resulted in blue-, cyan- and yellow-green light emitting versions. Despite the great utility of GFP, however, other fluorescent proteins with properties similar to or different from GFP would be useful in the art. Novel 20 fluorescent proteins result in possible new colors, or produce pH-dependent fluorescence. Other benefits of novel fluorescent proteins include fluorescence resonance energy transfer (FRET) possibilities based on new spectra and better suitability for larger excitation.

The prior art is deficient in novel fluorescent proteins wherein the DNA coding sequences are known. The present invention fulfills this long-standing need in the art.

25

## SUMMARY OF THE INVENTION

The present invention is directed to DNA sequences encoding fluorescent proteins selected from the group consisting of:  
(a) an isolated DNA from an organism from the Class Anthozoa which

encodes a fluorescent protein; (b) an isolated DNA which hybridizes to the isolated DNA of (a) and which encodes a fluorescent protein; and (c) an isolated DNA differing from the isolated DNAs of (a) and (b) in codon sequence due to the degeneracy of the genetic code and that 5 encodes a fluorescent protein. Preferably, the DNA is isolated from a non-bioluminescent organism from Class Anthozoa. More preferably, the DNA has the sequence selected from the group consisting of SEQ ID Nos. 55 and 57, and the fluorescent protein has the amino acid sequence selected from the group consisting of SEQ ID Nos. 56 and 58.

10 In another embodiment of the present invention, there is provided a vector capable of expressing the DNA of the present invention in a recombinant cell comprising said DNA and regulatory elements necessary for expression of the DNA in the cell. Preferably, the DNA encodes a fluorescent protein having the amino acid sequence 15 selected from the group consisting of SEQ ID Nos. 56 and 58.

15 In still another embodiment of the present invention, there is provided a host cell transfected with a vector of the present invention, such that the host cell expresses a fluorescent protein. Preferably, the cell is selected from the group consisting of bacterial 20 cells, mammalian cells, plant cells, insect cells and yeast cells.

25 The present invention is also directed to an isolated and purified fluorescent protein coded for by DNA selected from the group consisting of: (a) isolated DNA from an organism from Class Anthozoa which encodes a fluorescent protein; (b) isolated DNA which hybridizes to the isolated DNA of (a) and which encodes a fluorescent protein; and (c) isolated DNA differing from the isolated DNAs of (a) and (b) in codon sequence due to the degeneracy of the genetic code, and which encodes a fluorescent protein. Preferably, the protein has the amino

acid sequence selected from the group consisting of SEQ ID Nos. 56 and 58.

The present invention is also directed to a DNA sequence encoding a fluorescent protein selected from the group consisting of:

5 (a) an isolated DNA which encodes a fluorescent protein, wherein said DNA is from an organism from Class Anthozoa and wherein said organism does not exhibit bioluminescence; (b) an isolated DNA which hybridizes to isolated DNA of (a) and which encodes a fluorescent protein; and (c) an isolated DNA differing from the isolated DNAs of

10 (a) and (b) in codon sequence due to degeneracy of the genetic code and which encodes a fluorescent protein. Preferably, the organism is from Sub-class Zoantharia, Order Zoanthidea. More preferably, the organism is from Sub-order Brachycnemina. Even more preferably, the organism is from Family Zoanthidae, Genus Zoanthus. Most

15 particularly, the present invention is drawn to a novel fluorescent protein from *Zoanthus sp.*, zFP538.

The present invention is further directed to an amino acid sequence which can be used as a basis for designing an oligonucleotide probe for identification of a DNA encoding a fluorescent protein by means of hybridizaton, wherein the amino acid sequence is selected from the group consisting of SEQ ID Nos. 3, 5, 8, 11, 12, 14. Preferably, such an oligonucleotide has a nucleotide sequence selected from the group consisting of SEQ ID Nos. 4, 6, 7, 9, 10, 13, 15, 16.

Other and further aspects, features, and advantages of the present invention will be apparent from the following description of the presently preferred embodiments of the invention given for the purpose of disclosure.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the modified strategy of 3'-RACE used to isolate the target fragments. Sequences of the oligonucleotides used are shown in Table 2. Dp1 and Dp2 are the degenerate primers used in the first and second PCR, respectively (see Tables 3 and 4 for the sequences of degenerate primers). In the case of *Zoanthus sp.*, the first degenerate primer used was NGH (SEQ ID No. 4), and the second degenerate primer used was GEGa (SEQ ID No. 6).

10

Figure 2 shows the excitation and emission spectrum of the novel fluorescent protein from *Zoanthus sp.*, zFP538.

Figure 3 shows the excitation and emission spectrum of mutant zFP538, i.e., m128v.

Figure 4 shows transient expression of pYNFP m128v-N1 and pEYFP-N1 in 293 cells, respectively. pYNFP m128v-N1 (Figure 4B) shows as bright fluorescent intensity as pEYFP-N1 (Figure 4A) by fluorescent microscopy. The fluorescence showed is green (pseudocolor), however, the actual color should be yellow.

Figure 5 shows that fusion protein PKC- $\gamma$ -YNFP (M128V) translocated from cytosol to the plasma membrane when cells were treated with PMA (Phorbol 12-Myristate 13-Acetate). Figure 5A shows the result from control (without the treatment) and Figure 5B shows the result from PMA-treated cells.

## DETAILED DESCRIPTION OF THE INVENTION

As used herein, the term "GFP" refers to the basic green fluorescent protein from *Aequorea victoria*, including prior art versions of GFP engineered to provide greater fluorescence or fluoresce in different colors. The sequence of *Aequorea victoria* GFP (SEQ ID No. 54) has been disclosed in Prasher et al., *Gene* 111 (1992), 229-33.

As used herein, the term "EGFP" refers to mutant variant of GFP having two amino acid substitutions: F64L and S65T (Heim et al., 10 *Nature* 373 (1995), 663-664). The term "humanized" refers to changes made to the GFP nucleic acid sequence to optimize the codons for expression of the protein in human cells (Yang et al., *Nucleic Acids Research* 24 (1996), 4592-4593).

As used herein, the term "YFP" refers to yellow fluorescent protein, and the term "EYFP" refers to enhanced yellow fluorescent protein.

As used herein, the term "NFP" refers to novel fluorescent protein, and the term "YNFP" refers to yellow novel fluorescent protein. Specifically, "YNFP" refers to zFP538

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Maniatis, Fritsch & Sambrook, "Molecular Cloning: A Laboratory Manual" (1982); 20 "DNA Cloning: A Practical Approach," Volumes I and II (D.N. Glover ed. 1985); "Oligonucleotide Synthesis" (M.J. Gait ed. 1984); "Nucleic Acid Hybridization" (B.D. Hames & S.J. Higgins eds. (1985)); "Transcription and Translation" (B.D. Hames & S.J. Higgins eds. (1984)); "Animal Cell 25 Culture" (R.I. Freshney, ed. (1986)); "Immobilized Cells and Enzymes"

(IRL Press, (1986)); B. Perbal, "A Practical Guide To Molecular Cloning" (1984).

A "vector" is a replicon, such as plasmid, phage or cosmid, to which another DNA segment may be attached so as to bring about 5 the replication of the attached segment.

A "DNA molecule" refers to the polymeric form of deoxyribonucleotides (adenine, guanine, thymine, or cytosine) in either single stranded form or a double-stranded helix. This term refers only 10 to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary forms. Thus, this term includes double-stranded DNA found, inter alia, in linear DNA molecules (e.g., restriction fragments), viruses, plasmids, and chromosomes.

A DNA "coding sequence" is a DNA sequence which is transcribed and translated into a polypeptide in vivo when placed 15 under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic sequences, cDNA from eukaryotic mRNA, genomic DNA 20 sequences from eukaryotic (e.g., mammalian) DNA, and synthetic DNA sequences. A polyadenylation signal and transcription termination sequence may be located 3' to the coding sequence.

As used herein, the term "hybridization" refers to the process of association of two nucleic acid strands to form an 25 antiparallel duplex stabilized by means of hydrogen bonding between residues of the opposite nucleic acid strands.

The term "oligonucleotide" refers to a short (under 100 bases in length) nucleic acid molecule.

"DNA regulatory sequences", as used herein, are transcriptional and translational control sequences, such as promoters, enhancers, polyadenylation signals, terminators, and the like, that provide for and/or regulate expression of a coding sequence in a host cell.

A "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site, as well as protein binding domains responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain "TATA" boxes and "CAT" boxes. Various promoters, including inducible promoters, may be used to drive the various vectors of the present invention.

As used herein, the terms "restriction endonucleases" and "restriction enzymes" refer to bacterial enzymes, each of which cut double-stranded DNA at or near a specific nucleotide sequence.

A cell has been "transformed" or "transfected" by exogenous or heterologous DNA when such DNA has been introduced inside the cell. The transforming DNA may or may not be integrated (covalently linked) into the genome of the cell. In prokaryotes, yeast, and mammalian cells for example, the transforming DNA may be maintained on an episomal element such as a plasmid. With respect to eukaryotic cells, a stably transformed cell is one in which the transforming DNA has become integrated into a chromosome so that it

is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transforming DNA. A "clone" is a population of 5 cells derived from a single cell or common ancestor by mitosis. A "cell line" is a clone of a primary cell that is capable of stable growth *in vitro* for many generations.

A "heterologous" region of the DNA construct is an identifiable segment of DNA within a larger DNA molecule that is not 10 found in association with the larger molecule in nature. Thus, when the heterologous region encodes a mammalian gene, the gene will usually be flanked by DNA that does not flank the mammalian genomic DNA in the genome of the source organism. In another example, heterologous DNA includes coding sequence in a construct where 15 portions of genes from two different sources have been brought together so as to produce a fusion protein product. Allelic variations or naturally-occurring mutational events do not give rise to a heterologous region of DNA as defined herein.

As used herein, the term "reporter gene" refers to a coding 20 sequence attached to heterologous promoter or enhancer elements and whose product may be assayed easily and quantifiably when the construct is introduced into tissues or cells.

The amino acids described herein are preferred to be in the "L" isomeric form. The amino acid sequences are given in one-letter 25 code (A: alanine; C: cysteine; D: aspartic acid; E: glutamic acid; F: phenylalanine; G: glycine; H: histidine; I: isoleucine; K: lysine; L: leucine; M: methionine; N: asparagine; P: proline; Q: glutamine; R: arginine; S: serine; T: threonine; V: valine; W: tryptophane; Y: tyrosine; X: any residue). NH<sub>2</sub> refers to the free amino group present at the amino

terminus of a polypeptide. COOH refers to the free carboxy group present at the carboxy terminus of a polypeptide. In keeping with standard polypeptide nomenclature, *J Biol. Chem.*, 243 (1969), 3552-59 is used.

5       The present invention is directed to an isolated DNA selected from the group consisting of: (a) isolated DNA from an organism from the Class Anthozoa which encodes a fluorescent protein; (b) isolated DNA which hybridizes to isolated DNA of (a) and which encodes a fluorescent protein; and (c) isolated DNA differing 10 from the isolated DNAs of (a) and (b) in codon sequence due to the degeneracy of the genetic code, and which encodes a fluorescent protein. Preferably, the DNA has the sequence selected from the group consisting of SEQ ID Nos. 55 and 57, and the fluorescent protein has the amino acid sequence selected from the group consisting of SEQ ID 15 Nos. 56 and 58. More preferably, the DNA is zFP538 or M128V.

In another embodiment of the present invention, there is provided a vector capable of expressing the DNA of the present invention in a recombinant cell comprising said DNA and regulatory elements necessary for expression of the DNA in the cell. Specifically, 20 the DNA encodes a fluorescent protein having the amino acid sequence selected from the group consisting of SEQ ID Nos. 56 and 58. Preferably, the vector is constructed by amplifying the DNA and then inserting the amplified DNA into EGFP-N1 backbone, or by fusing different mouse ODC degradation domains such as d1, d2 and d376 to 25 the C-terminal of the DNA and then inserting the fusion DNA into EGFP-N1 backbone.

In still another embodiment of the present invention, there is provided a host cell transfected with the vector of the present invention, which expresses a fluorescent protein of the present

invention. Preferably, the cell is selected from the group consisting of bacterial cells, mammalian cells, plant cells, insect cells and yeast cells. A representative example of mammalian cell is HEK 293 cell and an example of bacterial cell is an *E. coli* cell.

5       The present invention is also directed to a DNA sequence encoding a fluorescent protein selected from the group consisting of: (a) an isolated DNA which encodes a fluorescent protein, wherein said DNA is from an organism from Class Anthozoa and wherein said organism does not exhibit bioluminescence; (b) an isolated DNA which 10 hybridizes to isolated DNA of (a) and which encodes a fluorescent protein; and (c) an isolated DNA differing from the isolated DNAs of (a) and (b) in codon sequence due to degeneracy of the genetic code and which encodes a fluorescent protein. Preferably, the organism is from Sub-class Zoantharia, Order Zoanthidea. More preferably, the 15 organism is from Sub-order Brachycnemina. Even more preferably, the organism is from Family Zoanthidae, Genus Zoanthus.

The present invention is also directed to an isolated and purified fluorescent protein coded for by DNA selected from the group consisting of: (a) an isolated protein encoded by a DNA which encodes 20 a fluorescent protein wherein said DNA is from an organism from Class Anthozoa and wherein said organism does not exhibit bioluminescence; (b) an isolated protein encoded by a DNA which hybridizes to isolated DNA of (a); and (c) an isolated protein encoded by a DNA differing from the isolated DNAs of (a) and (b) in codon sequence due to 25 degeneracy of the genetic code. Preferably, the isolated and purified fluorescent protein is zFP538.

The present invention is further directed to an amino acid sequence which can be used as a basis for designing an oligonucleotide probe for identification of a DNA encoding a fluorescent protein by

means of hybridization, wherein the amino acid sequence is selected from the group consisting of SEQ ID Nos. 3, 5, 8, 11, 12, 14. Preferably, such an oligonucleotide has a nucleotide sequence selected from the group consisting of SEQ ID Nos. 4, 6, 7, 9, 10, 13, 15, 16 and is used as 5 a primer in polymerase chain reaction. Alternatively, it can be used as a probe for hybridization screening of the cloned genomic or cDNA library.

The following examples are given for the purpose of illustrating various embodiments of the invention and are not meant to 10 limit the present invention in any fashion.

### EXAMPLE 1

15 Biological Material

Novel fluorescent proteins were identified from several genera of Anthozoa which do not exhibit any bioluminescence but have fluorescent color as observed under usual white light or ultraviolet light. Six species were chosen (see Table 1).

TABLE 1Anthozoa Species Used in This Study

Species	Area of Origination	Fluorescent Color
<i>Anemonia majano</i>	Western Pacific	bright green tentacle tips
<i>Clavularia</i> sp.	Western Pacific	bright green tentacles and oral disk
<i>Zoanthus</i> sp.	Western Pacific	green-yellow tentacles and oral disk
<i>Discosoma</i> sp. "red"	Western Pacific	orange-red spots oral disk
<i>Discosoma</i> sp. "striata"	Western Pacific	blue-green stripes on oral disk
<i>Discosoma</i> sp. "magenta"	Western Pacific	faintly purple oral disk
<i>Discosoma</i> sp. "green"	Western Pacific	green spots on oral disk
<i>Anemonia sulcata</i>	Mediterranean	purple tentacle tips

**EXAMPLE 2****cDNA Preparation**

Total RNA was isolated from the species of interest according to the protocol of Chomczynski and Sacchi (Chomczynski P., et al., *Anal. Biochem.* 162 (1987), 156-159). First-strand cDNA was synthesized starting with 1-3 µg of total RNA using SMART PCR cDNA synthesis kit (CLONTECH) according to the provided protocol with the only alteration being that the "cDNA synthesis primer" provided in the kit was replaced by the primer TN3 (5'- CGCAGTCGACCG(T)<sub>13</sub>, SEQ ID No. 1) (Table 2). Amplified cDNA samples were then prepared as described in the protocol provided except the two primers used for PCR were the TS primer (5'-AAGCAGTGGTATCAACGCAGAGT, SEQ ID No. 2 ) (Table 2) and the TN3 primer (Table 2), both in 0.1 µM concentration. Twenty to twenty-five PCR cycles were performed to amplify a cDNA sample. The amplified cDNA was diluted 20-fold in water and 1 µl of this dilution was used in subsequent procedures.

**TABLE 2**Oligos Used in cDNA Synthesis and RACE

5 TN3: 5'-CGCAGTCGACCG(T)<sub>13</sub>  
(SEQ ID No. 1)

T7-TN3: 5'-GTAATACGACTCACTATAAGGCCGCAGTCGACCG(T)<sub>13</sub>  
(SEQ ID No. 17)

10 TS-primer: 5'-AAGCAGTGGTATCAACGCAGAGT  
(SEQ ID No. 2)

15 T7-TS:  
5'-GTAATACGACTCACTATAAGGCAAGCAGTGGTATCAACGCAGAGT  
(SEQ ID No. 18)

T7: 5'-GTAATACGACTCACTATAAGGC  
(SEQ ID No. 19)

20 TS-oligo 5'-AAGCAGTGGTATCAACGCAGAGTACGCrGrGrG  
(SEQ ID No. 53)

25

**EXAMPLE 3****Oligo Design**

To isolate fragments of novel fluorescent protein cDNAs,  
5 PCR using degenerate primers was performed. Degenerate primers  
were designed to match the sequence of the mRNAs in regions that  
were predicted to be the most invariant in the family of fluorescent  
proteins. Four such stretches were chosen (Table 3) and variants of  
degenerate primers were designed. All such primers were directed to  
10 the 3'-end of mRNA. All oligos were gel-purified before use. Table 2  
shows the oligos used in cDNA synthesis and RACE.

**TABLE 3**

Key Amino Acid Stretches and Corresponding Degenerate Primers Used  
for Isolation of Fluorescent Proteins

5

Stretch Position according to A. victoria GFP (7)	Amino Acid Sequence of the Key Stretch	Degenerated Primer Name and Sequence
20-25	GXVN GH (SEQ ID No. 3)	NGH: 5'- GA(C,T) GGC TGC GT(A,T,G,C) AA(T,C) GG(A,T,G) CA (SEQ ID No. 4)
31-35	GE GEG (SEQ ID No. 5)	GEGa: 5'- GTT ACA GGT GA(A,G) GG(A,C) GA(A,G) GG (SEQ ID No. 6)
	GEG NG (SEQ ID No. 8)	GEGb: 5'- GTT ACA GGT GA(A,G) GG(T,G) GA(A,G) GG (SEQ ID No. 7)
		GNGa: 5'- GTT ACA GGT GA(A,G) GG(A,C) AA(C,T) GG (SEQ ID No. 9)
		GNGb: 5'- GTT ACA GGT GA(A,G) GG(T,G) AA(C,T) GG (SEQ ID No. 10)
127-131	GMNFP (SEQ ID No. 11)	NFP: 5' TTC CA(C,T) GGT
	GVNFP (SEQ ID No. 12)	(G,A)TG AA(C,T) TT(C,T) CC (SEQ ID NO. 13)
134-137	GPVM (SEQ ID No. 14)	PVMa: 5' CCT GCC (G,A)A(C,T) GGT CC(A,T,G,C) GT(A,C) ATG (SEQ ID NO. 15)
		PVMb: 5' CCT GCC (G,A)A(C,T) GGT CC(A,T,G,C) GT(G,T) ATG (SEQ ID NO. 16)

**EXAMPLE 4****Isolation of 3'-cDNA Fragments of nEPs**

The modified strategy of 3'-RACE was used to isolate the target fragments (see Figure 1). The RACE strategy involved two consecutive PCR steps. The first PCR step involved a first degenerate primer (Table 4) and the T7-TN3 primer (SEQ ID No. 17) which has a 3' portion identical to the TN3 primer used for cDNA synthesis (for sequence of T7-TN3, Table 2). The reason for substituting the longer T7-TN3 primer in this PCR step was that background amplification which occurred when using the shorter TN3 primer was suppressed effectively, particularly when the T7-TN3 primer was used at a low concentration (0.1  $\mu$ M) (Frohman et al., (1998) *PNAS USA*, 85, 8998-9002). The second PCR step involved the TN3 primer (SEQ ID No. 1, Table 2) and a second degenerate primer (Table 4).

**TABLE 4**

Combinations of Degenerate Primers for First and Second PCR Resulting in Specific Amplification of 3'-Fragments of nFP cDNA

Species	First Degenerate Primer	Second Degenerate Primer
<i>Anemonia majano</i>	NGH (SEQ ID No. 4)	GNGb (SEQ ID No. 10)
<i>Clavularia</i> sp.	NGH (SEQ ID No. 4)	GEGa (SEQ ID No. 6)
<i>Zoanthus</i> sp.	NGH (SEQ ID No. 4)	GEGa (SEQ ID No. 6)
<i>Discosoma</i> sp. "red"	NGH (SEQ ID No. 4)	GEGa (SEQ ID No. 6), NFP (SEQ ID No. 13) or PVMb (SEQ ID No. 16)
<i>Discosoma striata</i>	NGH (SEQ ID No. 4)	NFP (SEQ ID No. 13)
<i>Anemonia sulcata</i>	NGH (SEQ ID No. 4)	GEGa (SEQ ID No. 6) or NFP (SEQ ID No. 13)

5

The first PCR reaction was performed as follows: 1 µl of 20-fold dilution of the amplified cDNA sample was added into the reaction mixture containing 1X Advantage KlenTaq Polymerase Mix with provided buffer (CLONTECH), 200 µM dNTPs, 0.3 µM of first degenerate

primer (Table 4) and 0.1  $\mu$ M of T7-TN3 (SEQ ID No. 17) primer in a total volume of 20  $\mu$ l. The cycling profile was (Hybaid OmniGene Thermocycler, tube control mode): 1 cycle for 95°C, 10 sec.; 55°C, 1 min.; 72°C, 40 sec; 24 cycles for 95°C, 10 sec.; 62°C, 30 sec.; 72°C, 40 sec. The reaction was then diluted 20-fold in water and 1  $\mu$ l of this dilution was added to a second PCR reaction, which contained 1X Advantage KlenTaq Polymerase Mix with the buffer provided by the manufacturer (CLONTECH), 200  $\mu$ M dNTPs, 0.3  $\mu$ M of the second degenerate primer (Table 4) and 0.1  $\mu$ M of TN3 primer. The cycling 10 profile was (Hybaid OmniGene Thermocycler, tube control mode): 1 cycle for 95°C, 10 sec.; 55°C (for GEG/GNG or PVM) or 52°C (for NFP), 1 min.; 72°C, 40 sec; 13 cycles for 95°C, 10sec.; 62°C (for GEG/GNG or PVM) or 58°C (for NFP), 30 sec.; 72°C, 40 sec. The product of PCR was cloned into PCR-Script vector (Stratagene) according to the 15 manufacturer's protocol.

Different combinations of degenerate primers were tried in the first and second PCR reactions on the DNA from each species until a combination of primers was found that resulted in specific amplification--meaning that a pronounced band of expected size 20 (about 650-800 bp for NGH and GEG/GNG and 350-500 bp for NFP and PVM--sometimes accompanied by a few minor bands) was detected on agarose gel after two PCR reactions. The primer combinations of choice for different species of the Class Anthozoa are listed in Table 4. Some other primer combinations also resulted in amplification of 25 fragments of correct size, but the sequence of these fragments showed no homology to the other fluorescent proteins identified or to *Aequorea victoria* GFP.

EXAMPLE 5Obtaining Full-Length cDNA Copies

Upon sequencing the obtained 3'-fragments of novel fluorescent protein cDNAs, two nested 5'-directed primers were synthesized for cDNA (Table 5), and the 5' ends of the cDNAs were then amplified using two consecutive PCRs. In the next PCR reaction, the novel approach of "step-out PCR" was used to suppress background amplification. The step-out reaction mixture contained 1x Advantage KlenTaq Polymerase Mix using buffer provided by the manufacturer (CLONTECH), 200 µM dNTPs, 0.2 µM of the first gene-specific primer (see Table 5), 0.02 µM of the T7-TS primer (SEQ ID No. 18), 0.1 µM of T7 primer (SEQ ID No. 19) and 1 µl of the 20-fold dilution of the amplified cDNA sample in a total volume of 20 µl. The cycling profile was (Hybaid OmniGene Thermocycler, tube control mode): 23-27 cycles for 95°C, 10 sec.; 60°C, 30 sec.; 72°C, 40 sec. The product of amplification was diluted 50-fold in water and one µl of this dilution was added to the second (nested) PCR. The reaction contained 1X Advantage KlenTaq Polymerase Mix with provided buffer (CLONTECH), 200 µM dNTPs, 0.2 µM of the second gene-specific primer and 0.1 µM of TS primer (SEQ ID No. 2) in a total volume of 20 µl. The cycling profile was (Hybaid OmniGene Thermocycler, tube control mode): 12 cycles for 95°C, 10 sec.; 60°C, 30 sec.; 72°C, 40 sec. The product of amplification was then cloned into pAtlas vector (CLONTECH) according to the manufacturer's protocol.

**TABLE 5**Gene-Specific Primers Used for 5'-RACE

Species	First Primer	Second (Nested) Primer
Anemonia majano	5'-GAAATAGTCAGGCATACTGGT (SEQ ID No. 20)	5'-GTCAGGCATAC TGGTAGGAT (SEQ ID No. 21)
Clavularia sp.	5'-CTTGAAATAGTCTGCTATATC (SEQ ID No. 22)	5'-TCTGCTATATC GTCTGGGT (SEQ ID No. 23)
Zoanthus sp.	5'- GTTCTTGAAATAGTCTACTATGT (SEQ ID No. 24)	5'-GTCTACTATGTCTT GAGGAT (SEQ ID No. 25)
Discosoma sp. "red"	5'-CAAGCAAATGGCAAAGGTC (SEQ ID No. 26)	5'-CGGTATTGTGGCC TTCGTA (SEQ ID No. 27)
Discosoma striata	5'-TTGTCTTCTTCTGCACAAC (SEQ ID No. 28)	5'-CTGCACAACGG GTCCAT (SEQ ID No. 29)
Anemonia sulcata	5'-CCTCTATCTTCATTCCTGC (SEQ ID No. 30)	5'-TATCTTCATTCCT GCGTAC (SEQ ID No. 31)
Discosoma sp. "magenta"	5'-TTCAGCACCCCATCACGAG (SEQ ID No. 32)	5'-ACGCTCAGAGCTG GGTTCC (SEQ ID No. 33)
Discosoma sp. "green"	5'-CCCTCAGCAATCCATCACGTT (SEQ ID No. 34)	5'-ATTATCTCAGTGGA TGGTTC (SEQ ID No. 35)

## EXAMPLE 6

### Expression of NEPs in *E.coli*

To prepare a DNA construct for novel fluorescent protein expression, two primers were synthesized for each cDNA: a 5'-directed "downstream" primer with the annealing site located in the 3'-UTR of the cDNA and a 3'-directed "upstream" primer corresponding to the site of translation start site (not including the first ATG codon) (Table 6). Primers with SEQ ID Nos. 41 and 42 were the primers used to prepare the zFP538 DNA. Both primers had 5'-heels coding for a site for a restriction endonuclease; in addition, the upstream primer was designed so as to allow the cloning of the PCR product into the pQE30 vector (Qiagen) in such a way that resulted in the fusion of reading frames of the vector-encoded 6xHis-tag and NFP. The PCR was performed as follows: 1 µl of the 20-fold dilution of the amplified cDNA sample was added to a mixture containing 1x Advantage KlenTaq Polymerase Mix with buffer provided by the manufacturer (CLONTECH), 200 µM dNTPs, 0.2 µM of upstream primer and 0.2 µM of downstream primer, in a final total volume of 20 µl. The cycling profile was (Hybaid OmniGene Thermocycler, tube control mode): 23-27 cycles for 95°C, 10 sec.; 60°C, 30 sec.; 72°C, 40 sec. The product of this amplification step was purified by phenol-chlorophorm extraction and ethanol precipitation and then cloned into pQE30 vector using restriction endonucleases corresponding to the primers' sequence according to standard protocols.

All plasmids were amplified in XL-1 blue *E. coli* and purified by plasmid DNA miniprep kits (CLONTECH). The recombinant clones were selected by colony color, and grown in 3 ml of LB medium (supplemented with 100 µg/ml of ampicillin) at 37°C overnight. 100 µl

of the overnight culture was transferred into 200 ml of fresh LB medium containing 100 µg/ml of ampicillin and grown at 37°C, 200 rpm up to OD<sub>600</sub> 0.6-0.7. 1 mM IPTG was then added to the culture and incubation was allowed to proceed at 37°C for another 16 hours. The 5 cells were harvested and recombinant protein, which incorporated 6x His tags on the N-terminus, was purified using TALON™ metal-affinity resin according to the manufacturer's protocol (CLONTECH).

**TABLE 6**

Primers Used to Obtain Full Coding Region of nEPs for Cloning into Expression Construct

Species	Upstream Primer	Downstream Primer
Anemonia majano	5' -acatggatccgcctttcaaaca agtttatc (SEQ ID No. 36) BamHI	5'-tagtactcgagcttattcgta tttcagtgaaatc (SEQ ID No. 37) XhoI
Clavularia sp.	L: 5'-acatggatccaacatttttga gaaacg (SEQ ID No. 38) BamHI S: 5'-acatggatccaaagctctaacc accatg (SEQ ID No. 39) BamHI	5'-tagtactcgagcaacacaa accctcagacaa (SEQ ID No. 40) XhoI
Zoanthus sp.	5'- acatggatccgcgtcagtcaaag cacgg (SEQ ID No. 41) BamHI	5'-tagtactcgagggttggaaactacat tcttatca (SEQ ID No. 42) XhoI
Discosoma sp. "red"	5'- acatggatccagggtttccaagaat gttatac (SEQ ID No. 43) BamHI	5'-tagtactcgaggaggccaagttc agccta (SEQ ID No. 44) XhoI
Discosoma striata	5'- acatggatccagggttgttccaagagtgt (SEQ ID No. 45) BamHI	5'-tagcgagctctatcatgcctc gtcacct (SEQ ID No. 46) SacI
Anemonia sulcata	5'- acatggatccgccttctttaaagaagact (SEQ ID No. 47) BamHI	5'-tagtactcgagtcctgggagc ggcttg (SEQ ID No. 48) XhoI
Discosoma sp. "magenta"	5'- acatggatccagggttgttccaagaatgtat (SEQ ID No. 49) BamHI	5'-tagtactcgaggccattacg ctaatac (SEQ ID No. 50) XhoI
Discosoma sp. "green"	5'-acatggatccagggtgcacttaagaagaaatg (SEQ ID No. 51)	5'-tagtactcgagattcggttaat gccttg (SEQ ID No. 52)

**EXAMPLE 7****Novel Fluorescent Proteins and cDNAs Encoding the Proteins**

One of the full-length cDNAs encoding novel fluorescent proteins is described herein (zFP538). The nucleic acid sequence and deduced amino acid sequence are SEQ ID Nos. 55 and 56, respectively. The spectral properties of zFP538 are listed in Table 7, and the emission and excitation spectrum for zFP538 is shown in Figure 2.

10

**TABLE 7****Spectral Properties of the Isolated zFP538**

Species:	Zoanthus sp.	Max. Extinction Coefficient:	20,200
nFP Name:	zFP538	Quantum Yield	0.42
Absorbance Max. (nm):	528	Relative Brightness:*	0.38
Emission Max. (nm):	538		

20

\*relative brightness is extinction coefficient multiplied by quantum yield divided by the same value for *A. victoria* GFP.

25

**EXAMPLE 8****Construction of zFP538 Mutant**

One mutant of zFP538 was generated, M128V. M128V was generated by introducing a wrong nucleotide in PCR during site-specific mutagenesis at position 65. One bright yellow colony was obtained, and the sequence of this clone was performed. It showed that this clone

contained wild type amino acid Lysine (K) at position 65, but had a substitution from Methionine (M) to Valine (V) at position 128 (numbering according to GFP). The nucleic acid sequence of M128V is shown in SEQ ID No. 57, and the deduced amino acid sequence is shown 5 in SEQ ID No. 58.

Further investigations showed that M128V has spectral characteristics very similar to wild type protein zFP538 but folds much faster. Figure 3 shows the emission and excitation spectrum for M128V. Table 8 lists the spectral properties of M128V.

10

#### TABLE 8

##### Spectral Properties of the Isolated M128V

Species:	Zoanthus sp.	Max. Extinction Coefficient:	25,360
nFP Name:	M128V	Quantum Yield	0.43
Absorbance Max. (nm):	53	Relative Brightness:*	0.50
Emission Max. (nm):	540		

\*relative brightness is extinction coefficient multiplied by quantum yield divided by the same value for *A. victoria* GFP.

25

#### EXAMPLE 9

##### Construction and Functional Analysis of Vectors

Both wildtype (wt) and mutant zFP538 DNA were amplified via PCR and reconstructed to EGFP-N1 backbone. This vector has the 30

same multiple cloning sites as EGFP-N1. The nucleic acid sequence of pYNFPwt-N1 is shown in SEQ ID No. 59. The nucleic acid sequence of pYNFPM128V-N1 is shown in SEQ ID No. 60. Both pYNFPwt and pYNFPW128V keep the same multiple cloning sites as EGFP-N1.

5           Functional test of the generated vectors was performed by transient transfection in 293 cells. After 24-hour expression, pYNFPwt, pYNFPM128V and EYFP were compared side by side: pYNFPwt showed less fluorescent intensity than EYFP (data not shown); however, pYNFPM128V showed as bright fluorescent intensity as EYFP by  
10 fluorescent microscopy (Figures 4A and 4B).

#### EXAMPLE 10

15 Generation of Destabilized zFP538 Vectors as Transcription Reporters

By using the same technology for destabilized EGFP, destabilized zFP538 vectors were constructed by fusing different mouse ODC degradation domains such as d1 and d2 to the C-terminal of zFP538. The d1 version of destabilized YNFP has three E to A mutations  
20 within MODC degradation domain compared to d2 version. Vectors pYNFPM128V-MODCd1 and pYNFPM128V-MODCd2 were constructed in EGFP-N1 backbone.

25

#### EXAMPLE 11

Functional Analysis of Destabilized zFP538

Functional test of the destabilized zFP538 was performed by transient transfection in 293 cells. After 24-hour expression, the

fluorescent intensity was decreased gradually from d2 and d1 because of the fusion with different mouse ODC degradation domains. After 4-hour treatment with protein synthesis inhibitor cycloheximide, d2 fluorescent intensity did not change very much; however, d1 5 fluorescent intensity decreased further 50% of its original intensity. The half-life of d1 is around 4 hours.

M128V has fast folding and bright fluorescent intensity, which makes it useful for number of applications. Some fusion proteins were tested such as PKC-gamma-YNFP (M128V). PKC-gamma was 10 observed to translocate from cytosol to the plasma membrane when cells were treated with PMA (Phorbol 12-Myristate 13-Acetate) (Figures 5A and 5B).

15

## EXAMPLE 12

### Construction and Functional Test for Humanized M128V

Humanized M128V was generated, and then placed into the pEGFP-N1 backbone. This vector has the same mutiple cloning sites as 20 pEGFP-N1. Construction of C1 and pEGFP is in the process.

Any patents or publications mentioned in this specification are indicative of the levels of those skilled in the art to which the invention pertains. These patents and publications are incorporated by 25 reference to the same extent as if each individual publication was specifically and individually indicated to be incorporated by reference.

One skilled in the art will appreciate readily that the present invention is adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those objects and ends inherent

therein. The present examples, along with the methods, procedures, treatments, molecules, and specific compounds described herein, are presently representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the invention. Changes 5 to the methods and compounds, and other uses, will occur to those skilled in the art and are encompassed within the spirit of the invention as defined by the scope of the claims.

**WHAT IS CLAIMED IS:**

1. A DNA sequence encoding a fluorescent protein selected from the group consisting of:

5 (a) an isolated DNA which encodes a fluorescent protein, wherein said DNA is from an organism from a Class Anthozoa and wherein said organism does not exhibit bioluminescence;

(b) an isolated DNA which hybridizes to isolated DNA of (a) above and which encodes a fluorescent protein; and

10 (c) an isolated DNA differing from the isolated DNAs of (a) and (b) above in codon sequence due to degeneracy of the genetic code and which encodes a fluorescent protein.

15 2. The DNA sequence of claim 1, wherein said organism is from Sub-class Zoantharia.

3. The DNA sequence of claim 2, wherein said organism 20 is from Order Zoanthidea.

4. The DNA sequence of claim 3, wherein said organism is from Sub-order Brachycnemina.

25

5. The DNA sequence of claim 4, wherein said organism is from Family Zoanthidae.

6. The DNA sequence of claim 5, wherein said organism is from Genus Zoanthus.

5 7. A DNA sequence encoding a fluorescent protein selected from the group consisting of:

(a) an isolated DNA which encodes a fluorescent protein having a nucleotide sequence selected from the group consisting of SEQ ID Nos. 55 and 57;

10 (b) an isolated DNA which hybridizes to isolated DNA of (a) above and which encodes a fluorescent protein; and

(c) an isolated DNA differing from the isolated DNAs of (a) and (b) above in codon sequence due to degeneracy of the genetic code, and which encodes a fluorescent protein.

15

8. The DNA sequence of claim 7, wherein said DNA encodes a fluorescent protein having an amino acid sequence selected from the group consisting of SEQ ID Nos. 56 and 58.

20

9. The DNA sequence of claim 7, wherein said DNA is non-humanized or humanized DNA.

25

10. The DNA sequence of claim 7, wherein said DNA is zFP538 or M128V.

11. A vector capable of expressing the DNA sequence of claim 1 in a recombinant cell, wherein said vector comprising said DNA and regulatory elements necessary for expression of the DNA in the cell.

5

12. The vector of claim 11, wherein said DNA encodes a fluorescent protein having the amino acid sequence selected from the group consisting of SEQ ID Nos. 56 and 58.

10

13. The vector of claim 11, wherein said vector is constructed by amplifying said DNA and then inserting the amplified DNA into EGFP-N1 backbone.

15

14. The vector of claim 13, wherein said DNA is non-humanized or humanized DNA.

20

15. The vector of claim 13, wherein said DNA is zFP538 or M128V.

25

16. The vector of claim 11, wherein said vector is constructed by fusing different mouse ODC degradation domains to the C-terminal of said DNA and then inserting the fusion DNA into EGFP-N1 backbone.

17. The vector of claim 16, wherein said mouse ODC degradation domains are selected from the group consisting of d1, d2 and d376.

5

18. The vector of claim 16, wherein said DNA is non-humanized or humanized DNA.

10

19. The vector of claim 16, wherein said DNA is zFP538 or M128V.

15

20. A host cell transfected with the vector of claim 11, wherein said cell is capable of expressing a fluorescent protein.

20

21. The host cell of claim 20, wherein said cell is selected from the group consisting of bacterial cells, mammalian cells, plant cell, yeast and insect cells.

22. The host cell of claim 21, wherein said mammalian cell is HEK 293 cell.

25

23. The host cell of claim 21, wherein said bacterial cell is an *E. coli* cell.

24. An isolated and purified fluorescent protein coded for by DNA selected from the group consisting of:

- (a) an isolated DNA which encodes a fluorescent protein 5 from an organism from Class Anthozoa, wherein said organism does not exhibit bioluminescence;
- (b) an isolated DNA which hybridizes to isolated DNA of (a) above and which encodes a fluorescent protein; and
- (c) an isolated DNA differing from the isolated DNAs of 10 (a) and (b) above in codon sequence due to degeneracy of the genetic code and which encodes a fluorescent protein.

25. The isolated and purified fluorescent protein of claim 15 24, wherein said organism is from Sub-class Zoantharia.

26. The isolated and purified fluorescent protein of claim 25, wherein said organism is from Order Zoanthidea.

20

27. The isolated and purified fluorescent protein of claim 26, wherein said organism is from Sub-order Brachycnemina.

25

28. The isolated and purified fluorescent protein of claim 27, wherein said organism is from Family Zoanthidae.

29. The isolated and purified fluorescent protein of claim 28, wherein said organism is from Genus Zoanthus.

5 30. An isolated and purified fluorescent protein coded for by DNA selected from the group consisting of:

(a) isolated DNA which encodes a fluorescent protein having an amino acid sequence selected from the group consisting of SEQ ID Nos. 56 and 58;

10 (b) isolated DNA which hybridizes to isolated DNA of (a) above and which encodes a fluorescent protein; and

(c) isolated DNA differing from said isolated DNAs of (a) and (b) above in codon sequence due to degeneracy of the genetic code and which encodes a fluorescent protein.

15

31. The isolated and purified fluorescent protein of claim 30, wherein said protein is zFP538.

20

32. An amino acid sequence which can be used as a basis for designing an oligonucleotide probe for identification of a DNA encoding a fluorescent protein by means of hybridizaton, wherein said sequence is selected from the group consisting of SEQ ID Nos. 3, 5, 8, 25 11, 12, 14.

33. The amino acid sequence of claim 22, wherein said oligonucleotide has a nucleotide sequence selected from the group consisting of SEQ ID Nos. 4, 6, 7, 9, 10, 13, 15, 16

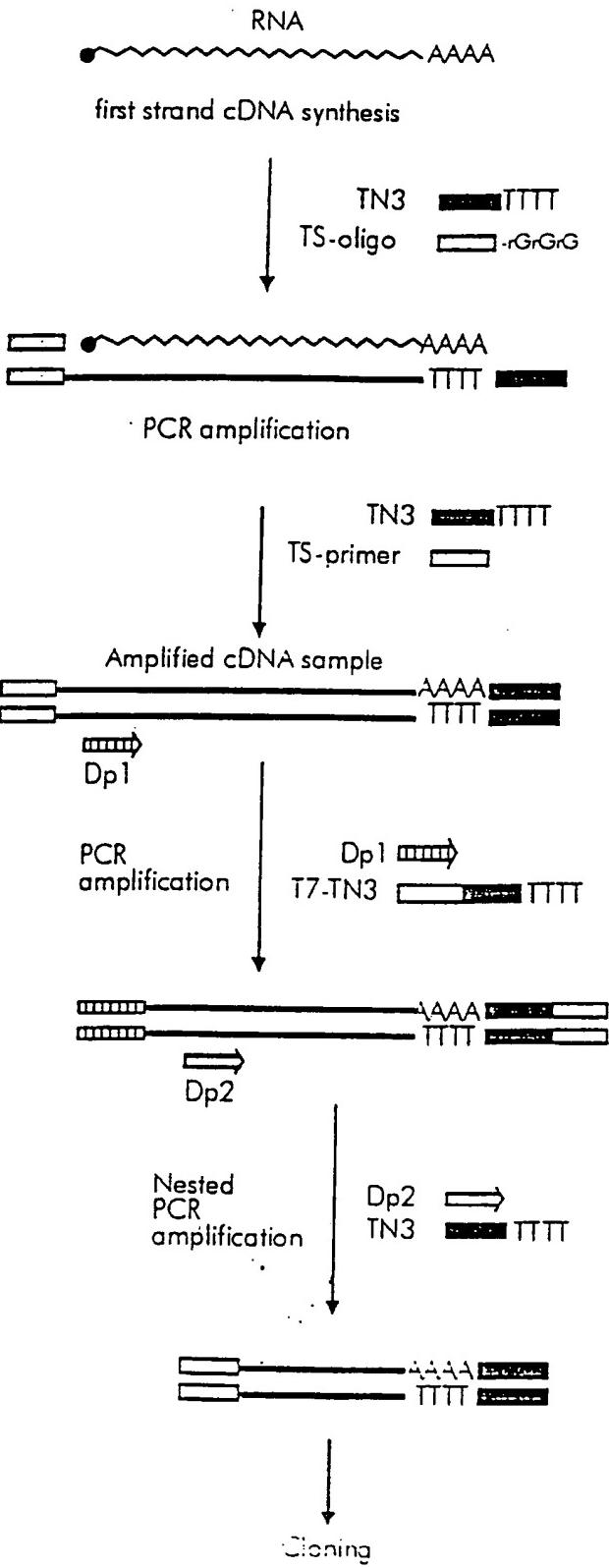


Figure 1

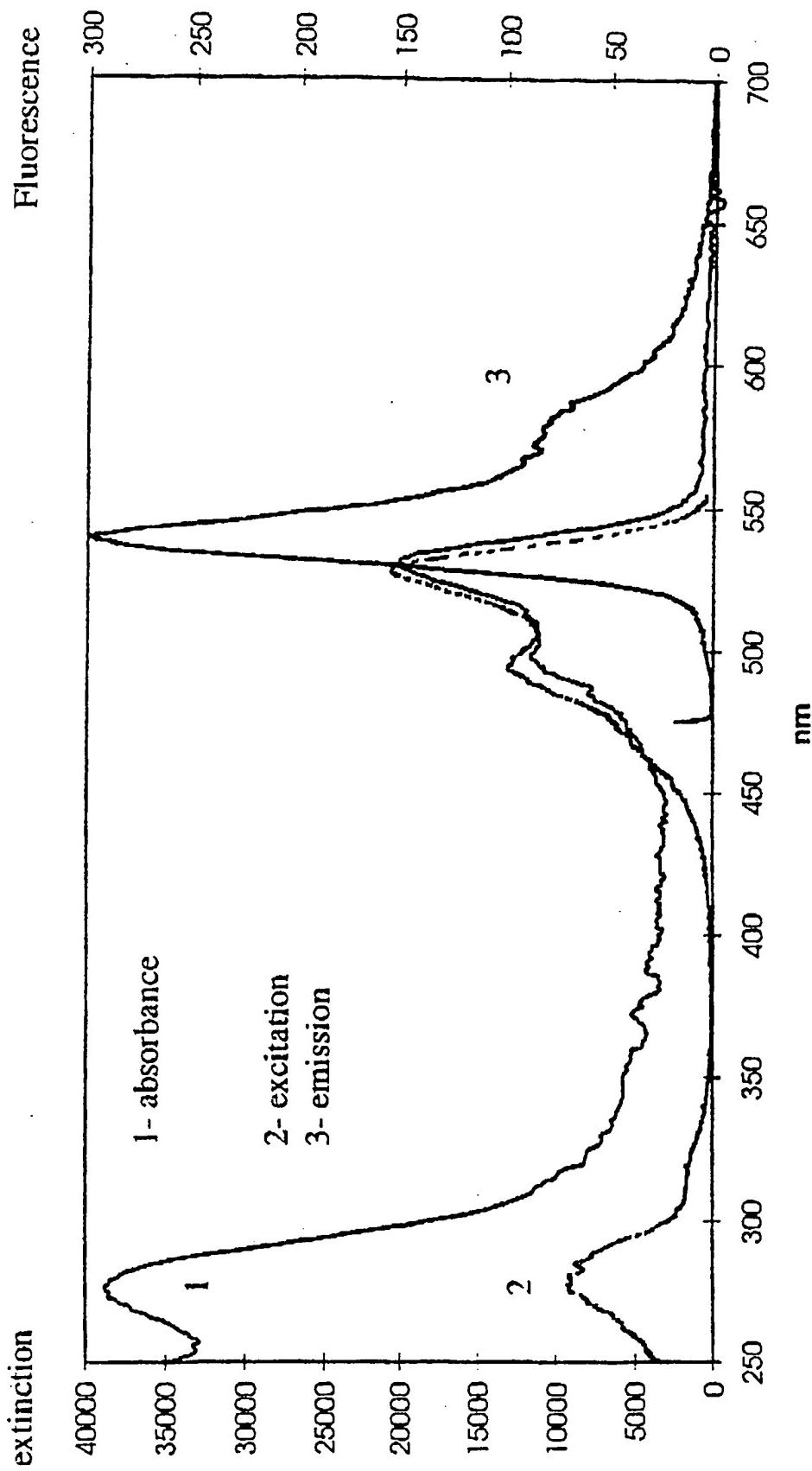


Figure 2

NFP4 -  $\lambda_{128\text{V}}^{\text{max}}$   
Spectrum

Date:  
20.08.99

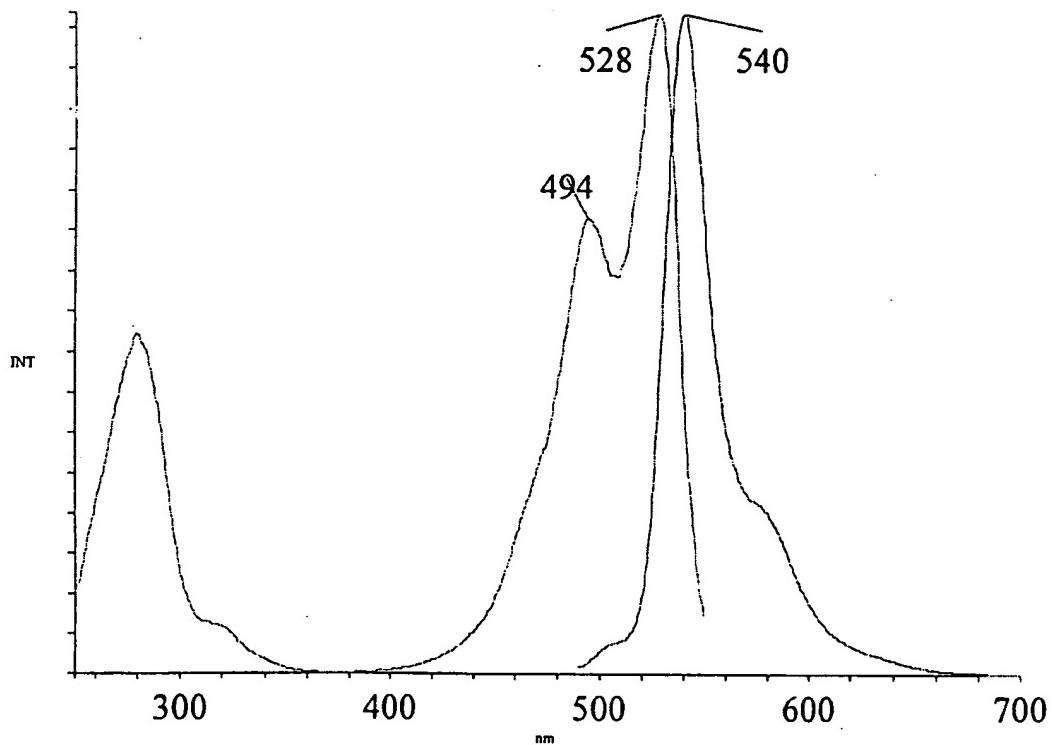


Figure 3

A. pEYFP-N1



B. pYNFP-N1 mut m128v

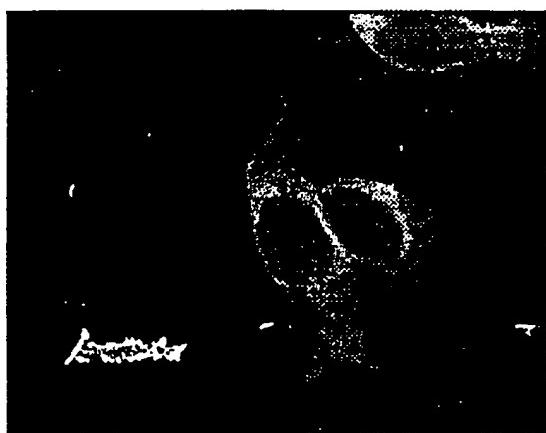


pseudocolor green. Should be yellow.

Figure 4

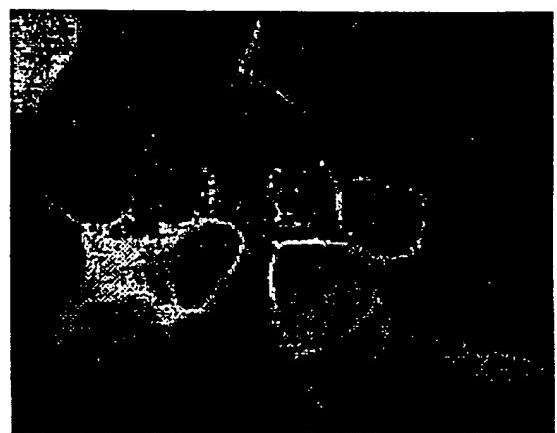
## PKCr-YFP translocation

A.



control

B.



P.M.A.

Figure 5

## SEQUENCE LISTING

<110> Lukyanov, Sergey A.  
Labas, Yuliia A.  
Matz, Mikhail V.  
5 Fradkov, Arcady F.  
Fang, Yu  
Tan, Wenyan

<120> Fluorescent proteins from non-bioluminescent  
species of Class Anthozoa, genes encoding such  
10 proteins and uses thereof

<130> D6196D4/PCT  
<141> 1999-12-09  
<150> 09/210,330  
<151> 1998-12-11

15 <160> 60

<210> 1  
<211> 25  
<212> DNA

20 <213> artificial sequence  
<220>  
<221> primer\_bind  
<223> primer TN3 used in cDNA synthesis and RACE  
<400> 1

25 cgcagtcgac cgttttttt ttttt 25

<210> 2  
<211> 23  
<212> DNA

30 <213> artificial sequence  
<220>  
<221> primer\_bind  
<223> primer TS used in cDNA synthesis and RACE  
<400> 2

35 aagcagtggatcaacgcag agt 23

<210> 3

<211> 6  
<212> PRT  
<213> *Aequorea victoria*  
<220>  
5 <222> 21  
<223> amino acid sequence of a key stretch on which  
primer NGH is based; Xaa at position 21  
represents unknown  
<400> 3

10 Gly Xaa Val Asn Gly His  
5

<210> 4  
<211> 20  
15 <212> DNA  
<213> artificial sequence  
<220>  
<221> primer\_bind  
<222> 12  
20 <223> primer NGH used for isolation of fluorescent  
protein; n at position 12 represents any of the  
four bases  
<400> 4

gayggctgcg tnaayggdca 20  
25 <210> 5  
<211> 5  
<212> PRT  
<213> *Aequorea victoria*  
30 <220>  
<222> 31..35  
<223> amino acid sequence of a key stretch on which  
primers GEGa and GEGb are based  
<400> 5

35 Gly Glu Gly Glu Gly  
5

<210> 6  
<211> 20  
<212> DNA  
<213> artificial sequence  
5 <220>  
<221> primer\_bind  
<223> primer GEGA used for isolation of fluorescent protein  
<400> 6

10 gttacaggtg arggmarggg 20

<210> 7  
<211> 20  
<212> DNA  
15 <213> artificial sequence  
<220>  
<221> primer\_bind  
<223> primer GEGb used for isolation of fluorescent protein  
20 <400> 7

gttacaggtg arggkgarggg 20

<210> 8  
<211> 5  
25 <212> PRT  
<213> *Aequorea victoria*  
<220>  
<222> 31..35  
<223> amino acid sequence of a key stretch on which  
30 primers GNGa and GNGb are based  
<400> 8

Gly Glu Gly Asn Gly  
5

35 <210> 9  
<211> 20  
<212> DNA

<213> artificial sequence  
<220>  
<221> primer\_bind  
5 <223> primer GNGa used for isolation of fluorescent  
protein  
<400> 9

gttacaggtg arggmaaygg 20

10 <210> 10  
<211> 20  
<212> DNA  
<213> artificial sequence  
<220>  
15 <221> primer\_bind  
<223> primer GNGb used for isolation of fluorescent  
protein  
<400> 10

gttacaggtg arggkaaygg 20

20 <210> 11  
<211> 5  
<212> PRT  
<213> *Aequorea victoria*  
<220>  
25 <222> 127..131  
<223> amino acid sequence of a key stretch on which  
primer NFP is based  
<400> 11

Gly Met Asn Phe Pro

30 5  
  
<210> 12  
<211> 5  
<212> PRT  
35 <213> *Aequorea victoria*  
<220>  
<222> 127..131

<223> amino acid sequence of a key stretch on which  
primer NFP is based  
<400> 12

Gly Val Asn Phe Pro

5 5

<210> 13  
<211> 20  
<212> DNA  
10 <213> artificial sequence  
<220>  
<221> primer\_bind  
<223> primer NFP used for isolation of fluorescent  
protein  
15 <400> 13

ttccayggtr tgaayt<sup>y</sup>cc 20

<210> 14  
<211> 4  
20 <212> PRT  
<213> *Aequorea victoria*  
<220>  
<222> 134..137  
<223> amino acid sequence of a key stretch on which  
primers PVMa and PVMb are based  
25 <400> 14

Gly Pro Val Met

30 <210> 15  
<211> 21  
<212> DNA  
<213> artificial sequence  
<220>  
35 <221> primer\_bind  
<222> 15

<223> primer PVMa used for isolation of fluorescent protein; n at position 15 represents any of the four bases  
<400> 15

5 cctgccrayg gtccngtmat g 21

<210> 16  
<211> 21  
<212> DNA  
10 <213> artificial sequence  
<220>  
<221> primer\_bind  
<222> 15  
<223> primer PVMb used for isolation of fluorescent protein; n at position 15 represents any of the four bases  
15 <400> 16

cctgccrayg gtccngtkat g 21

20 <210> 17  
<211> 47  
<212> DNA  
<213> artificial sequence  
<220>  
25 <221> primer\_bind  
<223> primer T7-TN3 used in cDNA synthesis and RACE  
<400> 17

gtaatacgac tcactatagg gccgcagtcg accgttttt tttttt  
47

30 <210> 18  
<211> 45  
<212> DNA  
<213> artificial sequence  
35 <220>  
<221> primer\_bind  
<223> primer T7-TS used in cDNA synthesis and RACE

<400> 18

gtaatacgcac tcactatagg gcaaggcagtgtatcaacgc agagt  
45

5 <210> 19  
<211> 22  
<212> DNA  
<213> artificial sequence  
<220>  
10 <221> primer\_bind  
<223> primer T7 used in cDNA synthesis and RACE  
<400> 19

gtaatacgcac tcactatagg gc 22

15 <210> 20  
<211> 21  
<212> DNA  
<213> artificial sequence  
<220>  
20 <221> primer\_bind  
<223> gene-specific primer used for 5'-RACE for  
*Anemonia majano*  
<400> 20

gaaatagtca ggcatactgg t 21

25 <210> 21  
<211> 20  
<212> DNA  
<213> artificial sequence  
30 <220>  
<221> primer\_bind  
<223> gene-specific primer used for 5'-RACE for  
*Anemonia majano*  
<400> 21

35 gtcaggcata ctggtaggat 20

<210> 22  
<211> 21  
<212> DNA  
5 <213> artificial sequence  
<220>  
<221> primer\_bind  
<223> gene-specific primer used for 5'-RACE for  
*Clavularia sp.*  
10 <400> 22

c ttgaaaatag tctgctatat c 21

<210> 23  
<211> 19  
15 <212> DNA  
<213> artificial sequence  
<220>  
<221> primer\_bind  
<223> gene-specific primer used for 5'-RACE for  
20 <223> *Clavularia sp.*  
<400> 23

tctgctatat cgtctgggt 19

<210> 24  
25 <211> 23  
<212> DNA  
<213> artificial sequence  
<220>  
<221> primer\_bind  
30 <223> gene-specific primer used for 5'-RACE for  
*Zoanthus sp.*  
<400> 24

gttcttgaaa tagtctacta tgt 23

35 <210> 25

<211> 20  
<212> DNA  
<213> artificial sequence  
<220>  
5 <221> primer\_bind  
<223> gene-specific primer used for 5'-RACE for  
*Zoanthus sp.*  
<400> 25

gtctactatg tcttgaggat 20

10 <210> 26  
<211> 19  
<212> DNA  
<213> artificial sequence  
15 <220>  
<221> primer\_bind  
<223> gene-specific primer used for 5'-RACE for  
*Discosoma sp. "red"*  
<400> 26

20 caagcaaatg gcaaaggtc 19

<210> 27  
<211> 19  
<212> DNA  
25 <213> artificial sequence  
<220>  
<221> primer\_bind  
<223> gene-specific primer used for 5'-RACE for  
*Discosoma sp. "red"*  
30 <400> 27

cggatttgtc gccttcgta 19

<210> 28  
<211> 19  
35 <212> DNA  
<213> artificial sequence

<220>  
<221> primer\_bind  
<223> gene-specific primer used for 5'-RACE for  
*Discosoma striata*

5 <400> 28

ttgtcttctt ctgcacaac 19

<210> 29  
<211> 17  
10 <212> DNA  
<213> artificial sequence  
<220>  
<221> primer\_bind  
<223> gene-specific primer used for 5'-RACE for  
15 *Discosoma striata*  
<400> 29

ctgcacaacg ggtccat 17

<210> 30  
20 <211> 20  
<212> DNA  
<213> artificial sequence  
<220>  
<221> primer\_bind  
25 <223> gene-specific primer used for 5'-RACE for  
*Anemonia sulcata*  
<400> 30

cctcttatctt catttcctgc 20

30 <210> 31  
<211> 20  
<212> DNA  
<213> artificial sequence  
<220>  
35 <221> primer\_bind  
<223> gene-specific primer used for 5'-RACE for  
*Anemonia sulcata*

	<400>	31	
	tatcttcatt tcctgcgtac		20
	<210>	32	
5	<211>	19	
	<212>	DNA	
	<213>	artificial sequence	
	<220>		
	<221>	primer_bind	
10	<223>	gene-specific primer used for 5'-RACE for <i>Discosoma sp.</i> "magenta"	
	<400>	32	
	ttcagcaccc catcacgag		19
	<210>	33	
15	<211>	19	
	<212>	DNA	
	<213>	artificial sequence	
	<220>		
20	<221>	primer_bind	
	<223>	gene-specific primer used for 5'-RACE for <i>Discosoma sp.</i> "magenta"	
	<400>	33	
	acgctcagag ctgggttcc		19
25	<210>	34	
	<211>	22	
	<212>	DNA	
	<213>	artificial sequence	
30	<220>		
	<221>	primer_bind	
	<223>	gene-specific primer used for 5'-RACE for <i>Discosoma sp.</i> "green"	
	<400>	34	
35	ccctcagcaa tccatcacgt tc		22

<210> 35  
<211> 20  
<212> DNA  
<213> artificial sequence  
5 <220>  
<221> primer\_bind  
<223> gene-specific primer used for 5'-RACE for  
*Discosoma sp.* "green"  
<400> 35

10 attatctcag tggatggtc 20  
  
<210> 36  
<211> 31  
<212> DNA  
15 <213> artificial sequence  
<220>  
<221> primer\_bind  
<223> upstream primer used to obtain full coding region  
of nFPs from *Anemonia majano*  
20 <400> 36  
  
acatggatcc gcttttcaa acaagtttat c 31  
  
<210> 37  
<211> 34  
25 <212> DNA  
<213> artificial sequence  
<220>  
<221> primer\_bind  
<223> downstream primer used to obtain full coding  
30 region of nFPs from *Anemonia majano*  
<400> 37  
  
tagtactcga gcttattcgt atttcagtga aatc 34  
  
<210> 38  
35 <211> 29  
<212> DNA  
<213> artificial sequence

<220>  
<221> primer\_bind  
<223> upstream primer used to obtain full coding region  
of nFPs from *Clavularia sp.*

5 <400> 38

acatggatcc aacattttt tgagaaaacg 29

<210> 39  
<211> 28  
10 <212> DNA  
<213> artificial sequence  
<220>  
<221> primer\_bind  
<223> upstream primer used to obtain full coding region  
15 of nFPs from *Clavularia sp.*  
<400> 39

acatggatcc aaagctctaa ccaccatg 28

<210> 40  
20 <211> 31  
<212> DNA  
<213> artificial sequence  
<220>  
<221> primer\_bind  
25 <223> downstream primer used to obtain full coding  
region of nFPs from *Clavularia sp.*  
<400> 40

tagtactcga gcaacacaaa ccctcagaca a 31

30 <210> 41  
<211> 28  
<212> DNA  
<213> artificial sequence  
<220>  
35 <221> primer\_bind  
<223> upstream primer used to obtain full coding region

of nFPs from *Zoanthus sp.*

<400> 41

acatggatcc gctcagtc aa agcacggt 28

5 <210> 42  
<211> 32  
<212> DNA  
<213> artificial sequence  
<220>  
10 <221> primer\_bind  
<223> downstream primer used to obtain full coding  
region of nFPs from *Zoanthus sp.*  
<400> 42

tagtactcg a ggttggaact acattcttat ca 32  
15 <210> 43  
<211> 31  
<212> DNA  
<213> artificial sequence  
20 <220>  
<221> primer\_bind  
<223> upstream primer used to obtain full coding region  
of nFPs from *Discosoma sp.* "red"  
<400> 43

25 acatggatcc aggtcttcca agaatgttat c 31  
  
<210> 44  
<211> 29  
<212> DNA  
30 <213> artificial sequence  
<220>  
<221> primer\_bind  
<223> downstream primer used to obtain full coding  
region of nFPs from *Discosoma sp.* "red"  
35 <400> 44

tagtactcg a ggagccaa gt tcagcctta 29

<210> 45  
<211> 28  
<212> DNA  
5 <213> artificial sequence  
<220>  
<221> primer\_bind  
<223> upstream primer used to obtain full coding region  
of nFPs from *Discosoma striata*  
10 <400> 45  
  
acatggatcc agttggtcca agagtgtg 28  
  
<210> 46  
<211> 28  
15 <212> DNA  
<213> artificial sequence  
<220>  
<221> primer\_bind  
<223> downstream primer used to obtain full coding  
region of nFPs from *Discosoma striata*  
20 <400> 46  
  
tagcgagctc tatcatgcct cgtcacct 28  
  
<210> 47  
25 <211> 31  
<212> DNA  
<213> artificial sequence  
<220>  
<221> primer\_bind  
30 <223> upstream primer used to obtain full coding region  
of nFPs from *Anemonia sulcata*  
<400> 47  
  
acatggatcc gtttccttt taaagaagac t 31  
  
35 <210> 48  
<211> 28  
<212> DNA

<213> artificial sequence  
<220>  
<221> primer\_bind  
5 <223> downstream primer used to obtain full coding  
region of nFPs from *Anemonia sulcata*  
<400> 48

tagtactcga gtccttggga gcggcttg 28

10 <210> 49  
<211> 30  
<212> DNA  
<213> artificial sequence  
<220>  
15 <221> primer\_bind  
<223> upstream primer used to obtain full coding region  
of nFPs from *Discosoma sp. "magenta"*  
<400> 49

acatggatcc agttgttcca agaatgtgat 30

20 <210> 50  
<211> 26  
<212> DNA  
<213> artificial sequence  
<220>  
25 <221> primer\_bind  
<223> downstream primer used to obtain full coding  
region of nFPs from *Discosoma sp. "magenta"*  
<400> 50

tagtactcga ggccattacg ctaatc 26

30 <210> 51  
<211> 31  
<212> DNA  
<213> artificial sequence  
35 <220>  
<221> primer\_bind  
<223> upstream primer used to obtain full coding region

of nFPs from *Discosoma sp.* "green"

<400> 51

acatggatcc agtgcactta aagaagaaat g 31

5 <210> 52

<211> 29

<212> DNA

<213> artificial sequence

<220>

10 <221> primer\_bind

<223> downstream primer used to obtain full coding  
region of nFPs from *Discosoma sp.* "green"

<400> 52

tagtactcga gattcggttt aatgccttg 29

15

<210> 53

<211> 33

<212> DNA

<213> artificial sequence

20 <220>

<221> primer\_bind

<223> TS-oligo used in cDNA synthesis and RACE

<400> 53

aaggcagtgg atcaacgcag agtacgcrgr grg 33

25

<210> 54

<211> 238

<212> PRT

30 <213> *Aequorea victoria*

<220>

<223> amino acid sequence of GFP

<400> 54

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu

35

5

10

15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser

	20	25	30
	Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys		
	35	40	45
	Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu		
5	50	55	60
	Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro		
	65	70	75
	Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu		
	80	85	90
10	Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn		
	95	100	105
	Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val		
	110	115	120
	Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn		
15	125	130	135
	Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val		
	140	145	150
	Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe		
	155	160	165
20	Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp		
	170	175	180
	His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu		
	185	190	195
	Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp		
25	200	205	210
	Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr		
	215	220	225
	Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys		
	230	235	
30			
	<210>	55	
	<211>	865	
	<212>	DNA	
	<213>	Zoanthus sp.	
35	<220>		
	<221>	CDS	
	<223>	cDNA sequence of zFP538	
	<400>	55	

gagttgagtt tctcgacttc agttgtatca attttggggc atcaagcgat 50  
 ctattttcaa catggctcat tcaaaggcacg gtctaaaaga agaaatgaca 100  
 atgaaatacc acatggaagg gtgcgtcaac ggacataaat ttgtgatcac 150  
 gggcgaaggc attggatatac cgttcaaagg gaaacagact attaatctgt 200  
 5 gtgtgatcga agggggacca ttgccattt ccgaagacat attgtcagct 250  
 ggcttaagt acggagacag gatttcact gaatatcctc aagacatagt 300  
 agactatttc aagaactcgt gtcctgctgg atatacatgg ggcaggtctt 350  
 ttctcttga gnatggagca gtctgcataat gcaatgtaga tataacagtg 400  
 agtgtcaaag aaaactgcat ttatcataag agcatattta atggaatgaa 450  
 10 ttttcctgct gatggacctg tgatgaaaaa gatgacaact aactgggaag 500  
 catcctgcga gaagatcatg ccagtaccta agcagggat actgaaaggg 550  
 gatgtctcca tgtacccct tctgaaggat ggtggcggtt accggtgcca 600  
 gttcgacaca gtttacaaag caaagtctgt gccaaagtaag atgcccggagt 650  
 ggcacattcat ccagcataag ctcctccgtg aagaccgcag cgatgctaag 700  
 15 aatcagaagt ggcagctgac agagcatgct attgcattcc cttctgcctt 750  
 ggcctgataa gaatgttagtt ccaacatttt aatgcatgtg cttgtcaatt 800  
 attctgataa aaatgttagtt gagttgaaaa cagacaagta caaataaagc 850  
 acatgtaaaat cgtct 865

20	<210>	56	
	<211>	230	
	<212>	PRT	
	<213>	<i>Zoanthus sp.</i>	
	<220>		
25	<223>	amino acid sequence of zFP538	
	<400>	56	
	Met Ala His Ser Lys His Gly Leu Lys Glu Glu Met Thr Met Lys		
	5	10	15
	Tyr His Met Glu Gly Cys Val Asn Gly His Lys Phe Val Ile Thr		
30	20	25	30
	Gly Glu Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Thr Ile Asn		
	35	40	45
	Leu Cys Val Ile Glu Gly Pro Leu Pro Phe Ser Glu Asp Ile		
	50	55	60
35	Leu Ser Ala Gly Phe Lys Tyr Gly Asp Arg Ile Phe Thr Glu Tyr		
	65	70	75
	Pro Gln Asp Ile Val Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly		

	80	85	90
	Tyr Thr Trp Gly Ser Phe Leu Phe Glu Asp Gly Ala Val Cys Ile		
	95	100	105
	Cys Asn Val Asp Ile Thr Val Ser Val Lys Glu Asn Cys Ile Tyr		
5	110	115	120
	His Lys Ser Ile Phe Asn Gly Met Asn Phe Pro Ala Asp Gly Pro		
	125	130	135
	Val Met Lys Lys Met Thr Thr Asn Trp Glu Ala Ser Cys Glu Lys		
	140	145	150
10	Ile Met Pro Val Pro Lys Gln Gly Ile Leu Lys Gly Asp Val Ser		
	155	160	165
	Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg Tyr Arg Cys Gln Phe		
	170	175	180
	Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Ser Lys Met Pro Glu		
15	185	190	195
	Trp His Phe Ile Gln His Lys Leu Leu Arg Glu Asp Arg Ser Asp		
	200	205	210
	Ala Lys Asn Gln Lys Trp Gln Leu Thr Glu His Ala Ile Ala Phe		
	215	220	225
20	Pro Ser Ala Leu Ala		
	230		

	<210>	57
	<211>	693
25	<212>	DNA
	<213>	artificial sequence
	<220>	
	<223>	nucleic acid sequence of M128V
	<400>	57
30	gctcattcaa agcacggtct aaaagaagaa atgacaatga aataccacat	50
	ggaagggtgc gtcaacggac ataaaatttgt gatcacgggc gaaggcattg	100
	gatatccgtt caaaggaaaa cagactattta atctgtgtgt gatcgaaggg	150
	ggaccattgc cattttccga agacatattg tcagctggct ttaagtacgg	200
	agacaggatt ttcactgaat atcctcaaga catagtagac tatttcaaga	250
35	actcggtgcc tgctggatat acatggggca ggtctttct ctttgaggat	300
	ggagcagtct gcatatgcaa tgttagatata acagtgagtg tcaaagaaaa	350
	ctgcatttat cataagagca tatttaatgg agtgaattt cctgctgatg	400

gacctgtgat gaaaaagatg acaactaact gggaaagcatc ctgcgagaag 450  
 atcatgccag tacctaagca ggggatactg aaaggggatg tctccatgta 500  
 cctccttctg aaggatggtg ggcgttaccg gtgccagttc gacacagtt 550  
 acaaagcaaa gtctgtgcca agtaagatgc cgagtgccca cttcatccag 600  
 5 cataagctcc tccgtgaaga ccgcagcgat gctaagaatc agaagtggca 650  
 gctgacagag catgctattg cattcccttc tgccctggcc tga 693

	<210>	58
	<211>	230
10	<212>	PRT
	<213>	artificial sequence
	<220>	
	<223>	amino acid sequence of M128V
	<400>	58
15	Ala His Ser Lys His Gly Leu Lys Glu Glu Met Thr Met Lys Tyr	
	5	10
	15	
	His Met Glu Gly Cys Val Asn Gly His Lys Phe Val Ile Thr Gly	
	20	25
	30	
	Glu Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Thr Ile Asn Leu	
20	35	40
	45	
	Cys Val Ile Glu Gly Pro Leu Pro Phe Ser Glu Asp Ile Leu	
	50	55
	60	
	Ser Ala Gly Phe Lys Tyr Gly Asp Arg Ile Phe Thr Glu Tyr Pro	
	65	70
	75	
25	Gln Asp Ile Val Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr	
	80	85
	90	
	Thr Trp Gly Arg Ser Phe Leu Phe Glu Asp Gly Ala Val Cys Ile	
	95	100
	105	
	Cys Asn Val Asp Ile Thr Val Ser Val Lys Glu Asn Cys Ile Tyr	
30	110	115
	120	
	His Lys Ser Ile Phe Asn Gly Val Asn Phe Pro Ala Asp Gly Pro	
	125	130
	135	
	Val Met Lys Lys Met Thr Thr Asn Trp Glu Ala Ser Cys Glu Lys	
	140	145
	150	
35	Ile Met Pro Val Pro Lys Gln Gly Ile Leu Lys Gly Asp Val Ser	
	155	160
	165	
	Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg Tyr Arg Cys Gln Phe	

	170	175	180
	Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Ser Lys Met Pro Glu		
	185	190	195
	Trp His Phe Ile Gln His Lys Leu Leu Arg Glu Asp Arg Ser Asp		
5	200	205	210
	Ala Lys Asn Gln Lys Trp Gln Leu Thr Glu His Ala Ile Ala Phe		
	215	220	225
	Pro Ser Ala Leu Ala		
	230		
10			
	<210>            59		
	<211>          4700		
	<212>          DNA		
	<213>          artificial sequence		
15	<220>		
	<223>          nucleic acid sequence of pYNFPzFP538-N1		
	<400>          59		
	tagttattaa tagtaatcaa ttacgggtc attagttcat agccatata        50		
	tggagttccg cgttacataa cttacggtaa atggccgc tggctgaccg        100		
20	cccaacgacc cccgcccatt gacgtcaata atgacgtatg ttcccatagt        150		
	aacgccaata gggactttcc attgacgtca atgggtggag tatttacggt        200		
	aaactgccc cttggcagta catcaagtgt atcatatgcc aagtacgccc        250		
	cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta        300		
	catgacctta tgggactttc ctacttgca gtacatctac gtattagtca        350		
25	tcgctattac catggtgatg cggtttggc agtacatcaa tggcggtgga        400		
	tagcggttg actcacgggg atttccaagt ctccacccca ttgacgtcaa        450		
	tgggagttt ttttggcacc aaaatcaacg ggactttcca aaatgtcgta        500		
	acaactccgc cccattgacg caaatggcg gtaggcgtgt acggtggag        550		
	gtctatataa gcagagctgg tttagtgaac cgtcagatcc gctagcgcta        600		
30	ccggactcag atctcgagct caagttcga attctgcagt cgacggtacc        650		
	gcgggccccg gatccggta catggctcag tcaaagcacg gtctaaaaga        700		
	agaaatgaca atgaaatacc acatggaagg gtgcgtcaac ggacataaat        750		
	ttgtgatcac gggcgaaggc attggatatc cgttcaaagg gaaacagact        800		
	attaatctgt gtgtgatcga agggggacca ttgccatccc ccgaagacat        850		
35	attgtcagct ggcttaagt acggagacag gatttcact gaatatcctc        900		
	aagacatagt agactatttc aagaactcgt gtccctgctgg atatacatgg        950		
	ggcaggtctt ttctctttga ggatggagca gtctgcataat gcaatgtaga 1000		

tataacagtg agtgtcaaag aaaactgcattatcataag agcatattta 1050  
 atggaaatgaa ttccctcgat gatggacctg tgatgaaaaa gatgacaact 1100  
 aactggaaag catcctgcga gaagatcatg ccagtaccta agcagggat 1150  
 actgaaaagg gatgtctcca tgtaccttctt ctgaaggatg gtggccgtta 1200  
 5 cccgtgccag ttgcacacag ttacaaagc aaagtctgtg ccaagtaaga 1250  
 tgccggagtg gcacttcatc cagcataagc tcctccgtga agaccgcagc 1300  
 gatgctaaga atcagaagtgc gtagactgaca gagcatgcta ttgcattccc 1350  
 ttctgccttgc gcctgagcgg ccgcgactct agatcataat cagccataacc 1400  
 acattttagt aggtttact tgctttaaaa aacctccac acctccccct 1450  
 10 gaacctgaaa cataaaatga atgcaattgt tttgttaac ttgtttattt 1500  
 cagcttataa tggttacaaa taaagcaata gcatcacaaa ttccacaaat 1550  
 aaagcatttt tttcactgca ttctagttgt ggttgttcca aactcatcaa 1600  
 tgtatcttaa ggcgtaaatt gtaagcgtta atatttgtt aaaattcgcg 1650  
 ttaaattttt gttaaatcag ctcattttt aaccaatagg ccgaaatcgg 1700  
 15 caaaatccct tataaatcaa aagaatagac cgagataggg ttgagtggtt 1750  
 ttccagtttgc gacaagagt ccactattaa agaacgtgga ctccaacgtc 1800  
 aaagggcgaa aaaccgtcta tcagggcgat ggcccactac gtgaaccatc 1850  
 accctaatac agtttttgg ggtcgaggtg ccgtaaagca ctaaatcgg 1900  
 accctaaagg gagcccccgat ttttagagctt gacggggaaa gccggcgaac 1950  
 20 gtggcgagaa aggaaggaa gaaagcgaaa ggagcggcgct ctagggcgct 2000  
 ggcaagtgttgc gcggtcacgc tgcgcttaac caccacaccc gccgcgtt 2050  
 atgcgcccgtt acagggcgcg tcaggtggca ctttcgggg aaatgtgcgc 2100  
 ggaaccctta tttgtttatt tttctaaata cattcaaata tgtatccgt 2150  
 catgagacaa taaccctgat aaatgcttca ataatattga aaaaggaaga 2200  
 25 gtcctgagggc ggaaagaacc agctgtggaa tttgtgtcag ttaggggtgt 2250  
 gaaagtcccc aggctccccca gcaggcagaa gtatgcaaag catgcacatc 2300  
 aattagttagtgc caaccaggtg tggaaagtcc ccaggctccc cagcaggcag 2350  
 aagtatgcaa agcatgcacatc tcaatttagtc agcaaccata gtcccgcccc 2400  
 taactccgccc catcccgcccc ctaactccgc ccagttccgc ccattctccg 2450  
 30 ccccatggct gactaatttt ttttatttgcagaggccg aggccgcctc 2500  
 ggcctctgag ctattccaga agtagtgagg aggctttttt ggaggcctag 2550  
 gctttgcaa agatcgatca agagacagga tgaggatcgat ttgcgtatgt 2600  
 tgaacaagat ggattgcacg caggttctcc ggccgcttgg gtggagaggc 2650  
 tattcggcta tgactggca caacagacaa tcggctgctc tgatgccgccc 2700  
 35 gtgttccggc tgtcagcgca gggcgccccg gttcttttgc tcaagaccga 2750  
 cctgtccggcgtt gcccctgaatg aactgcaaga cgaggcagcg cggctatcgt 2800  
 ggctggccac gacggcggtt cttgcgcag ctgtgctcga cgttgtcact 2850  
 gaagcggaa gggactggct gctattggc gaagtgccgg ggcaggatct 2900

cctgtcatct cacccggctc ctgcccggaaa agtatccatc atggctgatg 2950  
caatgcggcg gctgcatacg cttgatccgg ctacacctgccc attcgaccac 3000  
caagcgaaac atcgcatcga gcgagcacgt actcgatgg aagccggct 3050  
tgtcgatcag gatgatctgg acgaagagca tcaggggctc gcgccagccg 3100  
5 aactgttcgc caggctcaag gcgagcatgc ccgacggcga ggatctcgtc 3150  
gtgacccatg gcgatgcctg cttgccgaat atcatggtgg aaaatggccg 3200  
ctttctgga ttcatcgact gtggccggct ggggtgtggcg gaccgctatc 3250  
aggacatagc gttggctacc cgtgatattg ctgaagagct tggcggcgaa 3300  
tgggctgacc gcttcctcgt gctttacggt atcgccgctc ccgattcgca 3350  
10 ggcgcacatcgcc ttcttatcgcc ttcttgacga gttcttctga gccccactct 3400  
gggggttcgaa atgaccgacc aagcgacgcc caacctgcca tcacgagatt 3450  
tcgattccac cgccgccttc tatgaaaggt tgggcttcgg aatcgtttc 3500  
cgggacgccc gctggatgtat cctccagcgc gggatctca tgctggagtt 3550  
cttcgcccac cctaggggaa ggctaactga aacacggaag gagacaatac 3600  
15 cggaaggaac ccgcgcstatg acggcaataa aaagacagaa taaaacgcac 3650  
ggtgttgggt cgtttgttca taaacgcggg gttcggtccc agggctggca 3700  
ctctgtcgat acccccacca gaccccaattt gggccaatac gccccgcgttt 3750  
cttccttttc cccacccac cccccaagtt cgggtgaagg cccaggggctc 3800  
gcagccaacg tcggggcgcc agggccctgcc atagcctcag gttactcata 3850  
20 tataacttttag attgatttaa aacttcattt ttaatttaaa aggatctagg 3900  
tgaagatcct ttttgataat ctcatgacca aaatccctta acgtgagttt 3950  
tcgttccact gagcgtcaga cccccgtagaa aagatcaaag gatcttctt 4000  
agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaaccac 4050  
cgctaccagc ggtggtttgc ttgcccggatc aagagctacc aactctttt 4100  
25 ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgtccttct 4150  
agtgtagccg tagttaggcc accacttcaa gaactctgtt gcacccgccta 4200  
cataccctcg tctgctaattt ctgttaccag tggctgctgc cagtggcgat 4250  
aagtcgtgtc ttaccgggtt ggactcaaga cgatagtttccggataaggc 4300  
gcagcggcgtc ggctgaacgg ggggttcgtc cacacagccc agcttggagc 4350  
30 gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc 4400  
gccacgccttc ccgaagggag aaaggcggac aggtatccgg taagcggcag 4450  
ggtcggaaaca ggagagcgcgca cgagggagct tccagggggaa aacgcctggt 4500  
atctttatag tcctgtcggg tttcgccacc tctgacttga gcgtcgattt 4550  
ttgtgtatgct cgtcagggggg gcggagccta tggaaaaacgc ccagcaacgc 4600  
35 ggcctttta cgggttcctgg ccttttgctg gccttttgc cacatgttct 4650  
ttcctgcgtt atccctgtat tctgtggata accgtattac cgccatgcatt 4700

<210> 60  
<211> 4700  
<212> DNA  
<213> artificial sequence  
5 <220>  
<223> nucleic acid sequence of pYNFPM128V-N1  
<400> 60  
tagttattaa tagtaatcaa ttacgggtc attagttcat agcccatata 50  
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg 100  
10 cccaaacgacc cccgcccatt gacgtcaata atgacgtatg ttcccatagt 150  
aacgccaata gggactttcc attgacgtca atgggtggag tatttacggt 200  
aaactgcccc cttggcagta catcaagtgt atcatatgcc aagtacgccc 250  
cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 300  
catgaccta tgggactttc ctacttggca gtacatctac gtattagtca 350  
15 tcgctattac catggtgatg cggtttggc agtacatcaa tgggcgtgga 400  
tagcggtttg actcacgggg atttccaagt ctccacccca ttgacgtcaa 450  
tgggagtttgcacc aaaaatcaacg ggactttcca aaatgtcgta 500  
acaactccgc cccattgacg caaatggcg gttaggcgtgt acgggtggag 550  
gtctatataa gcagagctgg ttttagtgaac cgtcagatcc gctagcgcta 600  
20 ccggactcag atctcgagct caagcttcga attctgcagt cgacggtacc 650  
gcggggcccg gatccggtagt catggctcag tcaaaggcacg gtctaaaaga 700  
agaaatgaca atgaaataacc acatggaagg gtgcgtcaac ggacataaaat 750  
ttgtgatcac gggcgaaggc attggatatac cgttcaaagg gaaacagact 800  
attaatctgt gtgtgatcga agggggacca ttgccatttt ccgaagacat 850  
25 attgtcagct ggcttaagt acggagacag gatttcact gaatatcctc 900  
aagacatagt agactattc aagaactcgt gtcctgctgg atatacatgg 950  
ggcaggtctt ttctcttga ggatggagca gtctgcataat gcaatgtaga 1000  
tataacagtg agtgtcaaag aaaaactgcatttataag agcatattta 1050  
atggagtgaa ttttcctgct gatggacctg tgatgaaaaa gatgacaact 1100  
30 aactggaaag catcctgcga gaagatcatg ccagtaccta agcagggat 1150  
actgaaaggg gatgtctcca tgtacctctt ctgaaggatg gtggccgtta 1200  
cccggtccag ttgcacacag tttacaaagc aaagtctgtg ccaagtaaga 1250  
tgccggagtgcacttcatc cagcataagc tcctccgtga agaccgcagc 1300  
gatgctaaga atcagaagtgcagactgacagcatgctta ttgcattccc 1350  
35 ttctgccttgcctgagcgg cccgcacttgcgttgcataat cagccataacc 1400  
acattttagt aggtttact tgctttaaaa aacctcccac acctccccct 1450  
gaacctgaaa cataaaatga atgcaattgt tggtttaac ttgtttattg 1500  
cagcttataa tggttacaaa taaagcaata gcatcacaaa tttcacaaat 1550

aaagcatttt tttcaactgca ttcttagttgt ggtttgcaca aactcatcaa 1600  
 tgtatcttaa ggcgtaaatt gtaagcgta atattttgtt aaaattcgcg 1650  
 ttaaaatttt gttaaatcag ctcattttt aaccaatagg ccgaaatcg 1700  
 caaaaatccct tataaatcaa aagaatagac cgagataggg ttgagtgtt 1750  
 5 ttccagtttgaacaagagt ccactattaa agaacgtgga ctccaacgac 1800  
 aaagggcgaa aaaccgtcta tcagggcgat ggcccactac gtgaaccatc 1850  
 accctaatacg agtttttgg ggtcgaggtg ccgtaaagca ctaaatcg 1900  
 accctaaagg gagcccccgat ttttagagctt gacggggaaa gccggcgaac 1950  
 gtggcgagaa aggaaggaa gaaagcgaaa ggagcggcgct 2000  
 10 ggcaagtgtatgcgtcacgc tgcgcttaac caccacaccc gccgcgctt 2050  
 atgcgccgct acagggcgat tcaggtggca ctttcgggg aaatgtgcgc 2100  
 ggaaccccta tttgtttatt tttctaaata cattcaaata tgtatccgct 2150  
 catgagacaa taaccctgat aaatgcttca ataatattga aaaaggaaga 2200  
 gtcctgaggc ggaaagaacc agctgtggaa tgtgtgtcag ttagggtgt 2250  
 15 gaaagtcccc aggctccccca gcaggcagaa gtatgcaaag catgcac 2300  
 aatttagtcag caaccaggtg tggaaagtcc ccaggctccc cagcaggcag 2350  
 aagtatgcaa agcatgcac tcatttcgtc agcaaccata gtcccgcccc 2400  
 taactccgccc catcccgccc ctaactccgc ccagttccgc ccattctccg 2450  
 cccatggct gactaatttt ttttatttgcagaggccg aggccgcctc 2500  
 20 ggcctctgag ctattccaga agtagtgagg aggctttttt ggaggcctag 2550  
 gctttgcaa agatcgatca agagacagga tgaggatcgt ttgcgtatgt 2600  
 tgaacaagat ggattgcacg caggttctcc ggccgcttgg gtggagaggc 2650  
 tattcggcta tgactggca caacagacaa tcggctgctc tgatgccgc 2700  
 gtgttccggc tgtcagcgca gggcgccccg gttcttttgcagaggccg 2750  
 25 cctgtccggc gcccgtaatg aactgcaaga cgaggcagcg cggctatcgt 2800  
 ggctggccac gacggcggtt cttgcgcag ctgtgctcga cgttgtcact 2850  
 gaagcggaa gggactggct gctattggc gaagtgcgg ggcaggatct 2900  
 cctgtcatct caccttgctc ctgcccggaa agtatccatc atggctgtatg 2950  
 caatgcggcg gctgcatacg cttgatccgg ctacctgcctt attcgaccac 3000  
 30 caagcgaaac atcgcatcga gcgagcacgt actcggatgg aagccggct 3050  
 tgtcgatcag gatgatctgg acgaagagca tcaggggctc ggcgcggcc 3100  
 aactgttcgc caggctcaag gcgagcatgc ccgacggcga ggatctcg 3150  
 gtgaccatg gcgatgcctg cttgccaaat atcatggtgg aaaatggccg 3200  
 ctttcttggat ttcatcgact gtggccggct ggggtgtggcg gaccgctatc 3250  
 35 aggacatagc gttggcttacc cgtgatatttgc gttttacggatcgt tggcgccgaa 3300  
 tgggctgacc gcttcctcgat gctttacggatcgt atcgccgcctc ccgattcgca 3350  
 gcgcatcgcc ttcttatcgcc ttcttgacga gttcttctga gcgggactct 3400  
 ggggttcgaa atgaccgacc aagcgacgccc caacctgcca tcacgagatt 3450

tcgattccac cgccgccttc tatgaaaggt tgggcttcgg aatcgtttc 3500  
cgggacgccc gctggatgat cctccagcgc gggatctca tgctggagtt 3550  
cttcgcccac cctaggggga ggctaactga aacacggaag gagacaatac 3600  
cggaaggaac ccgcgctatg acggcaataa aaagacagaa taaaacgcac 3650  
5 ggtgttgggt cgtttgttca taaacgcggg gttcggtccc agggctggca 3700  
ctctgtcgat acccccaccga gaccccatcg gggccaatac gcccgcgttt 3750  
cttcctttc cccacccccac cccccaagtt cgggtgaagg cccagggctc 3800  
gcagccaaacg tcggggcggc aggccctgcc atagcctcag gttactcata 3850  
tatacttttag attgatttaa aacttcattt ttaatttaaa aggatctagg 3900  
10 tgaagatcct ttttataat ctcatgacca aaatccctta acgtgagttt 3950  
tcgttccact gagcgtcaga ccccgtagaa aagatcaaag gatcttcttg 4000  
agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac 4050  
cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactctttt 4100  
ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgtccttct 4150  
15 agttagccg tagttaggcc accacttcaa gaactctgta gcaccgccta 4200  
catacctcgc tctgctaattc ctgttaccag tggctgctgc cagtggcgat 4250  
aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac cggataaggc 4300  
gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc 4350  
gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc 4400  
20 gccacgcttc ccgaaggggag aaaggcggac aggtatccgg taagcggcag 4450  
ggtcggaaaca ggagagcgcgca cgagggagct tccagggggaa aacgcctgg 4500  
atctttatag tcctgtcggg tttcgccacc tctgacttga gcgtcgattt 4550  
ttgtgatgct cgtcagggggg gcggagccta tggaaaaacg ccagcaacgc 4600  
ggcctttta cggttcctgg cttttgctg gcctttgct cacatgttct 4650  
25 ttccctgcgtt atcccctgat tctgtggata accgtattac cgccatgcat 4700

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/29472

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(7) :C07K 14/435; C12N 1/00, 1/15, 1/21, 5/10, 15/12, 15/63  
 US CL :Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/320.1, 252.3, 252.33, 325, 410, 254.11, 348, 369, 69.1; 530/350, 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	MATZ et al. Fluorescent proteins from nonbioluminescent Anthozoa species. Nature Biotechnology. October 1999, Volume 17, No. 10, pages 969-973, see entire document.	1-33
X, P	DE 197 18 640 A1 (WIEDENMANN) 22 July 1999, (22.07.99), see entire document.	24-25, 30
A	US 5,491,084 A (CHALFIE et al) 13 February 1996 (13.02.96).	24-25, 30
X	ANDERLUH et al. Cloning, sequencing , and expression of equinatoxin II. Biochemical and Biophysical Research Communications. 1996, Volume 220, No. 2, pages 437-442, see entire document.	1-2, 7, 11, 20-21, 23-25, 30



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"B" earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

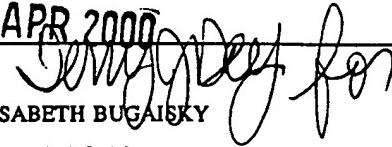
09 MARCH 2000

Date of mailing of the international search report

18 APR 2000

Name and mailing address of the ISA/US  
 Commissioner of Patents and Trademarks  
 Box PCT  
 Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer  
 GABRIELE ELISABETH BUGAISKY  


Telephone No. (703) 308-0196

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/29472

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MACEK et al. Intrinsic tryptophan fluorescence of equinatoxin II, a pore-forming polypeptide from the sea anemone <i>Actinia equina</i> L, monitors its interaction with lipid membranes. European Journal of Biochemistry. 1995, Volume 234, pages 329-335, entire document. Cited as "L" document because it establishes fluorescence of equinatoxin II.	24-25, 30
L		1-2, 7, 11, 20-21, 23

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/US99/29472

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims Nos.: 7-10, 12, 14-15, 30-33 because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

Since the sequence listing (CRF) submitted by applicant is defective, a sequence search could not be performed. Accordingly, claims 7-10, 12, 14-15 and 30-33 were searched only in-part based on a word search.
  
3.  Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
  
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**  

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/US99/29472

**A. CLASSIFICATION OF SUBJECT MATTER:**  
US CL :

435/320.1, 252.3, 252.33, 324, 410, 254.11, 348, 369, 69.1; 530/350; 536/23.5

**B. FIELDS SEARCHED**

Electronic data bases consulted (Name of data base and where practicable terms used):

Dialog files 155, 5, 434, 34, 358,28,44, 77 (Medline, Biosis, Scisearch, Derwent Biotech Abs., Oceanic Abs., Aquatic & Fish Abs., Dissertation Abs. Online, Conference Papers Index); STN-CAS files registry, CAPLUS; WEST files USPT, Derwent WPI

search terms: fluoresc?, bioluminesc?, protein?, polypeptide?, gene#, anthozo?, zoanth?, brachycnem? coral?, cnidar?, anemon?, zfp538, mahskhglik/sqsp, vngh/sqep, gegeg/sqep, gegng/sqep, gmnfp/sqep, gvnfp/sqep, gpvn/sqep